

Probenahme

Statistische Grundlagen der Probenahme
bei zweifacher Aufteilung
nach zwei gleichberechtigten Gesichtspunkten

DIN
53 803
Teil 3

Sampling; statistical basis; cross classification

Échantillonnage; base statistique; classification double

Inhalt

	Seite		Seite
1 Grundlagen	2	7.2 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit und minimalen Gesamtkosten	13
2 Begriffe	2	7.3 Stichprobenaufbau bei vorgegebenen Gesamtkosten	13
2.1 Begriffe für die Probenahme	2	7.4 Stichprobenaufbau bei vorgegebenen Gesamtkosten und kleinster erreichbarer Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit	13
2.2 Statistische Begriffe	2	8 Programmablaufpläne für die Berechnung des Stichprobenaufbaues	13
2.3 Weitere Begriffe	2	9 Probenahme bei reduzierbarer Aufteilung	13
3 Anordnungsschema der Einzelwerte bei zweifacher Aufteilung	2	Anhang A Beispiele aus der Textiltechnik	22
4 Statistische Auswertung bei zweifacher Aufteilung nach zwei gleichberechtigten Gesichtspunkten; Modell mit Wechselwirkung	2	A.1 Aufteilung nach zwei gleichberechtigten Gesichtspunkten; Modell mit Wechselwirkung; Durchführung der Varianzanalyse; Bestimmung von Schätzwerten für die Varianzen; Testen von Hypothesen über die Modellparameter; Vertrauensbereich für den Mittelwert der Gesamtheit; Luftdurchlässigkeit eines Papiermaschinenfilzes	22
4.1 Modellvorstellung	2	A.2 Stichprobenaufbau bei vorgegebener relativer Weite für den Vertrauensbereich der Gesamtheit; Modell ohne Wechselwirkung; varianzbestimmte kostenminimale Stichprobe; Luftdurchlässigkeit eines Papiermaschinenfilzes	24
4.2 Auswertung	2	A.3 Stichprobenaufbau bei vorgegebenen Gesamtkosten und kleinster erreichbarer Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; Modell ohne Wechselwirkung; kostenbestimmte varianzminimale Stichprobe; Luftdurchlässigkeit eines Papiermaschinenfilzes	25
4.3 Schätzwerte für die Modellparameter	3	A.4 Stichprobenaufbau bei vorgegebener relativer Weite für den Vertrauensbereich der Gesamtheit; Modell ohne Wechselwirkung; varianzbestimmte Stichprobe; Knittererholungswinkel eines Gewebes	25
4.4 Testen von Hypothesen über die Modellparameter	5	A.5 Stichprobenaufbau bei vorgegebenen Gesamtkosten; Modell ohne Wechselwirkung; kostenbestimmte Stichprobe; Knittererholungswinkel eines Gewebes	26
4.5 Vertrauensbereiche für die Modellparameter	5	A.6 Stichprobenaufbau bei vorgegebener relativer Weite für den Vertrauensbereich der Gesamtheit; Modell mit Wechselwirkung; varianzbestimmte Stichprobe; Rundversuch zur Messung der Höchstzugkraft eines Garns	26
4.6 Sonderfall: Modell ohne Wechselwirkung	5	Anhang B Formelzeichen	28
5 Stichprobenaufbau beim Modell mit Wechselwirkung	6	Zitierte Normen und andere Unterlagen	30
5.1 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; varianzbestimmte Stichprobe	6	Weitere Normen und andere Unterlagen	30
5.2 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit und minimalen Gesamtkosten; varianzbestimmte kostenminimale Stichprobe	8	Erläuterungen	30
5.3 Stichprobenaufbau bei vorgegebenen Gesamtkosten; kostenbestimmte Stichprobe	8		
5.4 Stichprobenaufbau bei vorgegebenen Gesamtkosten und kleinster erreichbarer Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; kostenbestimmte varianzminimale Stichprobe	9		
6 Stichprobenaufbau beim Modell ohne Wechselwirkung	9		
6.1 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; varianzbestimmte Stichprobe	9		
6.2 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit und minimalen Gesamtkosten; varianzbestimmte kostenminimale Stichprobe	10		
6.3 Stichprobenaufbau bei vorgegebenen Gesamtkosten; kostenbestimmte Stichprobe	11		
6.4 Stichprobenaufbau bei vorgegebenen Gesamtkosten und kleinster erreichbarer Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; kostenbestimmte varianzminimale Stichprobe	11		
7 Rundungsempfehlungen	12		
7.1 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit	12		

Fortsetzung Seite 2 bis 30

Normenausschuß Materialprüfung (NMP) im DIN Deutsches Institut für Normung e. V.

Ausschuß Qualitätssicherung und angewandte Statistik (AQS) im DIN

1 Grundlagen

In DIN 53 803 Teil 1 werden Verfahren zur Berechnung des Aufbaues und des Umfanges von Stichproben beschrieben, falls das zu prüfende Material in Gruppen zu unterteilen ist. Eine Unterteilung ist immer dann vorzunehmen, wenn Inhomogenitäten des Materials vorhanden sind oder vermutet werden.

Bei einer nach **einem** sachlichen Gesichtspunkt gewählten Einteilung in Gruppen ist eine einfache Aufteilung der Einheiten gegeben. Zur Auswertung werden Methoden der einfachen Varianzanalyse in DIN 53 803 Teil 1 herangezogen. Da in vielen Fällen eine einfache Aufteilung nicht ausreichend ist, wird man mehrfach aufteilen.

In dieser Norm und in DIN 53 803 Teil 4 wird der Aufbau und der Umfang von Stichproben behandelt, sofern man das zu untersuchende Material nach **zwei** sachlichen Gesichtspunkten unterteilt.

In dieser Norm wird die Unterteilung nach zwei **gleichberechtigten** Gesichtspunkten behandelt (Zweifache Varianzanalyse; Cross Classification), und zwar nur das Modell mit Zufallskomponenten. Zu Modellen mit systematischen Komponenten und gemischten Modellen wird auf die zitierten Unterlagen verwiesen. [1], [2]

Die Unterteilung in mehr als zwei Gruppen wird hier nicht behandelt. Falls eine weitergehende Unterteilung beabsichtigt ist, sollte man vorher überlegen, wie weit das sinnvoll ist und später befriedigend interpretiert werden kann. Zum Verständnis dieser Norm wird die Kenntnis von DIN 53 803 Teil 1 und Teil 2 vorausgesetzt.

2 Begriffe

2.1 Begriffe für die Probenahme

Begriffe für die Probenahme nach DIN 53 803 Teil 1, Ausgabe März 1979, Abschnitt 1.2.1 bis Abschnitt 1.2.11.

2.2 Statistische Begriffe

Statistische Begriffe nach DIN 53 803 Teil 1, Ausgabe März 1979, Abschnitt 1.3.1 bis Abschnitt 1.3.18. Gegenüber der dortigen Kennzeichnungsweise wird in dieser Norm jeder Einzelwert mit drei Indizes versehen. Sie kennzeichnen die Gruppe, zu der jeder Einzelwert gehört.

2.3 Weitere Begriffe

2.3.1 Zweifache Varianzanalyse

Bei der Gruppeneinteilung einer Gesamtheit nach zwei gleichberechtigten Gesichtspunkten, die zu einer Anordnung der Einzelwerte in Spalten und Zeilen führt, ist die zweifache Varianzanalyse ein statistisches Auswerteverfahren, das auf der Aufteilung der gesamten Summe der quadrierten Abweichungen (S.d.q.A.) in vier Anteile beruht:

- die S.d.q.A. zwischen den Zeilen (entsprechend der Unterteilung nach Gesichtspunkt 1),
- die S.d.q.A. zwischen den Spalten (entsprechend der Unterteilung nach Gesichtspunkt 2),
- die S.d.q.A. durch Wechselwirkung,
- die S.d.q.A. innerhalb der Zellen.

Daraus ergeben sich vier Varianzen, und zwar:

- die Varianz zwischen den Zeilen,
- die Varianz zwischen den Spalten,
- die Varianz durch Wechselwirkung,
- die Restvarianz.

3 Anordnungsschema der Einzelwerte bei zweifacher Aufteilung

Die Einzelwerte werden nach dem Schema in Tabelle 1 (linker Teil) zusammengestellt. Das Schema enthält n_1 Zeilen

($i = 1, 2, 3 \dots n_1$) entsprechend der Unterteilung nach Gesichtspunkt 1 und n_2 Spalten ($j = 1, 2, 3 \dots n_2$) entsprechend der Unterteilung nach Gesichtspunkt 2. An den Schnittpunkten von Zeilen und Spalten stehen in jeder Zelle (gekennzeichnet durch das Indexpaar ij) n_3 Einzelwerte x_{ijk} ($k = 1, 2, 3 \dots n_3$). Dabei wird in dieser Norm die Anzahl der Einzelwerte in jeder Zelle als gleich groß vorausgesetzt. Aus Gründen der Zweckmäßigkeit trägt man in dieses oder ein separates Schema sogleich die Quadrate x_{ijk}^2 der Einzelwerte in gleicher Weise ein.

Man bildet sodann innerhalb jeder Zelle die Zellensumme für die Einzelwerte

$$A_{ij} = \sum_{k=1}^{n_3} x_{ijk} \quad (1)$$

sowie die Quadrate der Einzelwerte

$$B_{ij} = \sum_{k=1}^{n_3} x_{ijk}^2 \quad (2)$$

Bis hierher ist der Gang der Rechnung unabhängig davon, ob man die Aufteilung nach zwei einander nachgeordneten oder nach zwei gleichberechtigten Gesichtspunkten wählt.

4 Statistische Auswertung bei zweifacher Aufteilung nach zwei gleichberechtigten Gesichtspunkten; Modell mit Wechselwirkung

4.1 Modellvorstellung

Bei der Aufteilung nach zwei gleichberechtigten Gesichtspunkten setzt sich jeder mögliche Einzelwert a_{ijk} der Gesamtheit aus mehreren Komponenten zusammen:

$$a_{ijk} = \mu + \xi_i + \eta_j + \alpha_{ij} + \varepsilon_{ijk} \quad (3)$$

Dabei ist

- μ der Mittelwert der Gesamtheit,
- ξ_i der Einfluß von Zeile i , d. h. von der Aufteilung nach Gesichtspunkt 1,
- η_j der Einfluß von Spalte j , d. h. von der Aufteilung nach Gesichtspunkt 2,
- α_{ij} die Wechselwirkung; das Wechselwirkungsglied ist allen Einzelwerten der Zelle ij gemeinsam, und es ist spezifisch für die Zelle ij ,
- ε_{ijk} der Zufallseinfluß beim Einzelwert k innerhalb der Zelle ij .

Für die Summanden in der Modellgleichung beim Modell mit Zufallskomponenten gilt folgendes:

- Die Summanden ξ_i , η_j und α_{ij} sind Zufallsgrößen, deren Erwartungswerte Null sind und deren Varianzen mit σ_ξ^2 , σ_η^2 und σ_α^2 bezeichnet werden.
- Die Varianzen innerhalb aller Zellen sollen gleich sein und werden mit σ_ε^2 (für alle ij) bezeichnet.
- Damit enthält das Modell mit Zufallskomponenten insgesamt die fünf Parameter μ , σ_ξ , σ_η , σ_α und σ_ε .

Für die Behandlung des Modells mit systematischen Komponenten und des gemischten Modells wird auf das Schrifttum verwiesen. [1]

Die Zufallsgrößen ξ_i , η_j , α_{ij} und ε_{ijk} sind normalverteilt.

4.2 Auswertung

Analog zur Modellgleichung wird jeder Meßwert x_{ijk} in folgende Anteile zerlegt:

$$x_{ijk} = \bar{x} \dots + (\bar{x}_{i \dots} - \bar{x} \dots) + (\bar{x}_{. j \dots} - \bar{x} \dots) + (\bar{x}_{ij \dots} - \bar{x}_{i \dots} - \bar{x}_{. j \dots} + \bar{x} \dots) + (x_{ijk} - \bar{x}_{ij \dots}) \quad (4)$$

Dabei ist

ein Einzelwert der Stichprobe:

$$x_{ijk}$$

der Gesamtmittelwert der Stichprobe:

$$\bar{x}_{...} = \frac{1}{n_1 \cdot n_2 \cdot n_3} \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \quad (5)$$

der Mittelwert der Zeile i :

$$\bar{x}_{i..} = \frac{1}{n_2 \cdot n_3} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \quad (6)$$

der Mittelwert der Spalte j :

$$\bar{x}_{.j.} = \frac{1}{n_1 \cdot n_3} \sum_{i=1}^{n_1} \sum_{k=1}^{n_3} x_{ijk} \quad (7)$$

der Mittelwert der Zelle ij :

$$\bar{x}_{ij.} = \frac{1}{n_3} \sum_{k=1}^{n_3} x_{ijk} \quad (8)$$

Ausgehend von dem Anordnungsschema im dick umrandeten Teil der Tabelle 1 für die Einzelwerte x_{ijk} (und die quadrierten Einzelwerte x_{ijk}^2) werden die folgenden auf dem rechten und unteren Rand eingetragenen Zwischensummen gebildet.

Summe aller Einzelwerte in der Zeile i :

$$\sum_{j=1}^{n_2} A_{ij} = \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \quad (9)$$

Summe aller quadrierten Einzelwerte in der Zeile i :

$$\sum_{j=1}^{n_2} B_{ij} = \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk}^2 \quad (10)$$

Summe aller Einzelwerte in der Spalte j :

$$\sum_{i=1}^{n_1} A_{ij} = \sum_{i=1}^{n_1} \sum_{k=1}^{n_3} x_{ijk} \quad (11)$$

Summe aller quadrierten Einzelwerte in der Spalte j :

$$\sum_{i=1}^{n_1} B_{ij} = \sum_{i=1}^{n_1} \sum_{k=1}^{n_3} x_{ijk}^2 \quad (12)$$

Quadrat der Summe aller Einzelwerte in der Zeile i :

$$\left(\sum_{j=1}^{n_2} A_{ij} \right)^2 = \left(\sum_{j=1}^{n_2} \sum_{k=1}^{n_3} x_{ijk} \right)^2 \quad (13)$$

Quadrat der Summe aller Einzelwerte in der Spalte j :

$$\left(\sum_{i=1}^{n_1} A_{ij} \right)^2 = \left(\sum_{i=1}^{n_1} \sum_{k=1}^{n_3} x_{ijk} \right)^2 \quad (14)$$

Zeilensumme der quadrierten Zellensummen:

$$\sum_{j=1}^{n_2} A_{ij}^2 = \sum_{j=1}^{n_2} \left(\sum_{k=1}^{n_3} x_{ijk} \right)^2 \quad (15)$$

Spaltensumme der quadrierten Zellensummen:

$$\sum_{i=1}^{n_1} A_{ij}^2 = \sum_{i=1}^{n_1} \left(\sum_{k=1}^{n_3} x_{ijk} \right)^2 \quad (16)$$

Diese für jede Zeile i und jede Spalte j vorliegenden vier Zwischensummen werden ihrerseits jeweils über alle Zeilen und Spalten zu folgenden Hilfsgrößen summiert:

Summe aller Einzelwerte:

$$A = \sum_i \sum_j A_{ij} = \sum_j \sum_i A_{ij} = \sum_i \sum_j \sum_k x_{ijk} \quad (17)$$

Summe aller quadrierten Einzelwerte:

$$B = \sum_i \sum_j B_{ij} = \sum_j \sum_i B_{ij} = \sum_i \sum_j \sum_k x_{ijk}^2 \quad (18)$$

Summe der Quadrate der Summe aller Einzelwerte innerhalb der Zeile i :

$$C = \sum_i \left(\sum_j A_{ij} \right)^2 = \sum_i \left(\sum_j \sum_k x_{ijk} \right)^2 \quad (19)$$

Summe der Quadrate der Summe aller Einzelwerte innerhalb der Spalte j :

$$D = \sum_j \left(\sum_i A_{ij} \right)^2 = \sum_j \left(\sum_i \sum_k x_{ijk} \right)^2 \quad (20)$$

Gesamtsumme der quadrierten Zellensummen:

$$E = \sum_i \sum_j A_{ij}^2 = \sum_j \sum_i A_{ij}^2 = \sum_i \sum_j \left(\sum_k x_{ijk} \right)^2 \quad (21)$$

Bei den Hilfsgrößen A , B und E gibt es zwei Rechenwege, nämlich die spaltenweise sowie die zeilenweise Addition der entsprechenden Zwischensummen. Es ist zweckmäßig, diese Beziehungen zur Rechenkontrolle zu benutzen. Die Kontrollen der Hilfsgrößen C und D sind erst später möglich.

Die Zwischensummen und Hilfsgrößen dürfen nicht gerundet werden (siehe DIN 53803 Teil 1, Ausgabe März 1979, Abschnitt 4, und DIN 53804 Teil 1, Ausgabe September 1981, Abschnitt 4.3).

4.3 Schätzwerte für die Modellparameter

Mit Hilfe der Summe der quadrierten Abweichungen Q und der zugehörigen Zahl der Freiheitsgrade f wird die Zerlegungstafel (siehe Tabelle 2) aufgestellt.

Die Summe Q_{total} läßt sich gemäß der Zerlegungstafel aus den Hilfsgrößen A und B errechnen; außerdem gilt jedoch die Beziehung

$$Q_{\text{total}} = Q_Z + Q_S + Q_{ZS} + Q_R \quad (22)$$

was zweckmäßig zur Kontrolle der Rechnung benutzt wird.

Analog gilt für die Zahl der Freiheitsgrade

$$f_{\text{total}} = n_1 \cdot n_2 \cdot n_3 - 1 = f_Z + f_S + f_{ZS} + f_R \quad (23)$$

wobei $n_1 \cdot n_2 \cdot n_3$ der gesamte Stichprobenumfang ist.

Als Ergebnis der Auswertung erhält man:

Schätzwert für μ ist

$$\hat{\mu} = x_{...} = \frac{A}{n_1 \cdot n_2 \cdot n_3} \quad (24)$$

Schätzwert für σ_{ξ}^2 ist

$$\hat{\sigma}_{\xi}^2 = \frac{s_Z^2 - s_{ZS}^2}{n_2 \cdot n_3} \quad (25)$$

Schätzwert für σ_{η}^2 ist

$$\hat{\sigma}_{\eta}^2 = \frac{s_S^2 - s_{ZS}^2}{n_1 \cdot n_3} \quad (26)$$

Schätzwert für σ_{α}^2 ist

$$\hat{\sigma}_{\alpha}^2 = \frac{s_{ZS}^2 - s_R^2}{n_3} \quad (27)$$



Tabelle 1. Anordnungsschema für die Einzelwerte und deren Quadrate; Zwischensummen und Hilfsgrößen für die zweifache Varianzanalyse

$i \backslash j$	1	2	bis	j	bis	n ₂	$\sum_j x_{ijk}$	$\sum_j \sum_k x_{ijk}^2$	$\left(\sum_j \sum_k x_{ijk}\right)^2$	$\sum_j \left(\sum_k x_{ijk}\right)^2$
1	x_{111} . . x_{11k} . . x_{11n3} A_{11}	x_{121} . . x_{12k} . . x_{12n3} A_{12}	bis	x_{1j1} . . x_{1jk} . . x_{1jn3} A_{1j}	bis	x_{1n21} . . x_{1n2k} . . x_{1n2n3} A_{1n2}	$\sum_j A_{1j}$	$\sum_j B_{1j}$	$\left(\sum_j A_{1j}\right)^2$	$\sum_j A_{1j}^2$
2	x_{211} . . x_{21k} . . x_{21n3} A_{21}	x_{221} . . x_{22k} . . x_{22n3} A_{22}	bis	x_{2j1} . . x_{2jk} . . x_{2jn3} A_{2j}	bis	x_{2n21} . . x_{2n2k} . . x_{2n2n3} A_{2n2}	$\sum_j A_{2j}$	$\sum_j B_{2j}$	$\left(\sum_j A_{2j}\right)^2$	$\sum_j A_{2j}^2$
bis	bis	bis		bis		bis	bis	bis	bis	bis
i	x_{i11} . . x_{i1k} . . x_{i1n3} A_{i1}	x_{i21} . . x_{i2k} . . x_{i2n3} A_{i2}	bis	x_{ij1} . . x_{ijk} . . x_{ijn3} A_{ij}	bis	x_{in21} . . x_{in2k} . . x_{in2n3} A_{in2}	$\sum_j A_{ij}$	$\sum_j B_{ij}$	$\left(\sum_j A_{ij}\right)^2$	$\sum_j A_{ij}^2$
bis	bis	bis		bis		bis	bis	bis	bis	bis
n ₁	x_{n111} . . x_{n11k} . . x_{n11n3} A_{n11}	x_{n121} . . x_{n12k} . . x_{n12n3} A_{n12}	bis	x_{n1j1} . . x_{n1jk} . . x_{n1jn3} A_{n1j}	bis	x_{n1n21} . . x_{n1n2k} . . x_{n1n2n3} A_{n1n2}	$\sum_j A_{n1j}$	$\sum_j B_{n1j}$	$\left(\sum_j A_{n1j}\right)^2$	$\sum_j A_{n1j}^2$
$\sum_j \sum_k x_{ijk}$	$\sum_j A_{1j}$	$\sum_j A_{12}$	bis	$\sum_j A_{1j}$	bis	$\sum_j A_{in2}$	A	↑	C	↑
$\sum_j \sum_k x_{ijk}^2$	$\sum_j B_{1j}$	$\sum_j B_{12}$	bis	$\sum_j B_{1j}$	bis	$\sum_j B_{in2}$	→	B		
$\sum_j \left(\sum_k x_{ijk}\right)^2$	$\left(\sum_j A_{1j}\right)^2$	$\left(\sum_j A_{12}\right)^2$	bis	$\left(\sum_j A_{1j}\right)^2$	bis	$\left(\sum_j A_{in2}\right)^2$	D			
$\sum_j \left(\sum_k x_{ijk}\right)^2$	$\sum_j A_{1j}^2$	$\sum_j A_{12}^2$	bis	$\sum_j A_{1j}^2$	bis	$\sum_j A_{in2}^2$	→			E

Tabelle 2. Zerlegungstafel für die zweifache Varianzanalyse beim Modell mit Zufallskomponenten

Schwankungen	Summe der quadrierten Abweichungen Q	Zahl der Freiheitsgrade f	Varianz $s^2 = \frac{Q}{f}$	Die Varianz ist Schätzwert für
zwischen den Zeilen	$Q_Z = n_2 \cdot n_3 \sum_{n_1} (\bar{x}_{i..} - \bar{x}_{...})^2 = \frac{C}{n_2 \cdot n_3} - \frac{A^2}{n_1 \cdot n_2 \cdot n_3}$	$f_Z = n_1 - 1$	s_Z^2	$\sigma_\epsilon^2 + n_3 \sigma_\alpha^2 + n_2 n_3 \sigma_\xi^2$
zwischen den Spalten	$Q_S = n_1 \cdot n_3 \sum_{n_2} (\bar{x}_{.j.} - \bar{x}_{...})^2 = \frac{D}{n_1 \cdot n_3} - \frac{A^2}{n_1 \cdot n_2 \cdot n_3}$	$f_S = n_2 - 1$	s_S^2	$\sigma_\epsilon^2 + n_3 \sigma_\alpha^2 + n_1 n_3 \sigma_\eta^2$
Wechselwirkung	$Q_{ZS} = n_3 \sum_{n_1} \sum_{n_2} (\bar{x}_{ij.} - \bar{x}_{i..} - \bar{x}_{.j.} + \bar{x}_{...})^2$ $= \frac{E}{n_3} - \frac{C}{n_2 \cdot n_3} - \frac{D}{n_1 \cdot n_3} + \frac{A^2}{n_1 \cdot n_2 \cdot n_3}$	$f_{ZS} = (n_1 - 1)(n_2 - 1)$	s_{ZS}^2	$\sigma_\epsilon^2 + n_3 \sigma_\alpha^2$
innerhalb der Zellen	$Q_R = \sum_{n_1} \sum_{n_2} \sum_{n_3} (x_{ijk} - \bar{x}_{ij.})^2 = B - \frac{E}{n_3}$	$f_R = n_1 n_2 (n_3 - 1)$	s_R^2	σ_ϵ^2
total	$Q_{\text{total}} = \sum_{n_1} \sum_{n_2} \sum_{n_3} (x_{ijk} - \bar{x}_{...})^2 = B - \frac{A^2}{n_1 \cdot n_2 \cdot n_3}$	$f_{\text{total}} = n_1 \cdot n_2 \cdot n_3 - 1$		

Schätzwert für σ_ϵ^2 ist

$$\hat{\sigma}_\epsilon^2 = s_R^2 \quad (28)$$

Falls einer oder mehrere der Schätzwerte $\hat{\sigma}_\xi^2$, $\hat{\sigma}_\eta^2$ und $\hat{\sigma}_\alpha^2$ negativ sind, werden die entsprechenden Varianzen durch Null geschätzt.

4.4 Testen von Hypothesen über die Modellparameter

Ausgehend von der Zerlegungstafel läßt sich mit dem F -Test prüfen, ob die Varianzen σ_ξ^2 , σ_η^2 und σ_α^2 als von Null verschieden anzusehen sind. Damit sind drei statistische Tests notwendig, die nach dem gleichen Schema ablaufen. Zunächst soll der Test der Hypothese über den Wechselwirkungseinfluß näher erläutert werden. Man stellt die Nullhypothese H_0 auf:

$$H_0: \sigma_\alpha^2 = 0, \text{ d. h. die Varianz der Wechselwirkung ist nicht von Null verschieden.}$$

Die Alternativhypothese H_1 lautet:

$$H_1: \sigma_\alpha^2 \neq 0, \text{ d. h. die Varianz der Wechselwirkung ist von Null verschieden.}$$

Die Nullhypothese wird mit dem Prüfwert s_{ZS}^2/s_R^2 geprüft, der mit dem Tabellenwert $F_{f_{ZS}; f_R; 1-\alpha}$ der F -Verteilung zum Signifikanzniveau α verglichen wird; die Tabellenwerte der F -Verteilung erhält man beispielsweise aus [1].

Ist $s_{ZS}^2/s_R^2 \leq F_{f_{ZS}; f_R; 1-\alpha}$, wird die Nullhypothese nicht verworfen und die Varianz σ_α^2 als nicht von Null verschieden angesehen.

Ist $s_{ZS}^2/s_R^2 > F_{f_{ZS}; f_R; 1-\alpha}$, wird die Nullhypothese zu Gunsten der Alternativhypothese verworfen, und die Varianz σ_α^2 wird als von Null verschieden betrachtet.

Der Test der Hypothese über den Zeileinfluß und der Test der Hypothese über den Spalteneinfluß laufen nach dem gleichen Schema ab. Alle drei Tests sind in der Tabelle 3 zusammengestellt.

4.5 Vertrauensbereiche für die Modellparameter

Der Vertrauensbereich für den Mittelwert der Gesamtheit μ zum Vertrauensniveau $1-\alpha$ ergibt sich näherungsweise zu

$$\bar{x}_{...} - t_{f; 1-\alpha/2} \cdot \sqrt{\frac{\hat{\sigma}_\xi^2}{n_1} + \frac{\hat{\sigma}_\eta^2}{n_2} + \frac{\hat{\sigma}_\alpha^2}{n_1 \cdot n_2} + \frac{\hat{\sigma}_\epsilon^2}{n_1 \cdot n_2 \cdot n_3}} \leq \mu \leq \bar{x}_{...} + t_{f; 1-\alpha/2} \cdot \sqrt{\frac{\hat{\sigma}_\xi^2}{n_1} + \frac{\hat{\sigma}_\eta^2}{n_2} + \frac{\hat{\sigma}_\alpha^2}{n_1 \cdot n_2} + \frac{\hat{\sigma}_\epsilon^2}{n_1 \cdot n_2 \cdot n_3}} \quad (29)$$

Der Freiheitsgrad für den Tabellenwert der t -Verteilung ist der kleinere der beiden Werte f_Z oder f_S . Wird eine der Varianzen σ_ξ^2 , σ_η^2 oder σ_α^2 als nicht von Null verschieden angesehen, entfällt in der Gleichung (29) der entsprechende Summand unter der Wurzel.

Der Vertrauensbereich für die Rest-Standardabweichung σ_ϵ ist

$$x_u \cdot s_R \leq \sigma_\epsilon \leq x_o \cdot s_R \quad (30)$$

wobei die Faktoren für x zum Freiheitsgrad f_R dem Schrifttum zu entnehmen sind. Vertrauensbereiche für die Standardabweichungen σ_ξ , σ_η und σ_α kann man nur näherungsweise angeben (siehe [1]).

4.6 Sonderfall: Modell ohne Wechselwirkung

Falls für ein bestimmtes Merkmal die Wechselwirkung bei der Zerlegung nach zwei gleichberechtigten Gesichtspunkten erfahrungsgemäß als nicht von Null verschieden anzusehen ist ($\alpha_{ij} = 0$ für alle i, j und daher $\sigma_\alpha^2 = 0$), genügt es oft, bei weiteren gleichartigen Untersuchungen nur eine Beobachtung je Zelle anzustellen ($n_3 = 1$).

In der Zerlegungstafel (Tabelle 2) verliert man zwar die Größen Q_R und f_R , da in diesem Falle $\sigma_\alpha^2 = 0$ und $n_3 = 1$ ist, jedoch erhält man aus der Summe Q_{ZS} und der Zahl der Freiheitsgrade f_{ZS} einen Schätzwert für σ_ϵ^2 .

Das Modell mit Zufallskomponenten hat also nur vier Modellparameter, die wie folgt geschätzt werden:

μ wird geschätzt durch

$$\hat{\mu} = \bar{x}_{...} = \bar{x} = \frac{A}{n_1 \cdot n_2} \quad (31)$$

σ_ξ^2 wird geschätzt durch

$$\hat{\sigma}_\xi^2 = \frac{s_Z^2 - s_{ZS}^2}{n_2} \quad (32)$$



Tabelle 3. Zusammenstellung der Tests für die verschiedenen Einflüsse zum Signifikanzniveau α ; Modell mit Wechselwirkung

Test für den Einfluß		
der Wechselwirkung $H_0: \sigma_\alpha^2 = 0; H_1: \sigma_\alpha^2 \neq 0$		
Prüfwert	s_{ZS}^2 / s_R^2	
Tabellenwert	$F_{f_{ZS}; f_R; 1-\alpha}$	
Vergleich	$s_{ZS}^2 / s_R^2 \leq F_{f_{ZS}; f_R; 1-\alpha}$	$s_{ZS}^2 / s_R^2 > F_{f_{ZS}; f_R; 1-\alpha}$
Ergebnis	$\sigma_\alpha^2 = 0$	$\sigma_\alpha^2 \neq 0$
der Zeilen $H_0: \sigma_\xi^2 = 0; H_1: \sigma_\xi^2 \neq 0$		
Prüfwert	s_Z^2 / s_{ZS}^2	
Tabellenwert	$F_{f_Z; f_{ZS}; 1-\alpha}$	
Vergleich	$s_Z^2 / s_{ZS}^2 \leq F_{f_Z; f_{ZS}; 1-\alpha}$	$s_Z^2 / s_{ZS}^2 > F_{f_Z; f_{ZS}; 1-\alpha}$
Ergebnis	$\sigma_\xi^2 = 0$	$\sigma_\xi^2 \neq 0$
der Spalten $H_0: \sigma_\eta^2 = 0; H_1: \sigma_\eta^2 \neq 0$		
Prüfwert	s_S^2 / s_{ZS}^2	
Tabellenwert	$F_{f_S; f_{ZS}; 1-\alpha}$	
Vergleich	$s_S^2 / s_{ZS}^2 \leq F_{f_S; f_{ZS}; 1-\alpha}$	$s_S^2 / s_{ZS}^2 > F_{f_S; f_{ZS}; 1-\alpha}$
Ergebnis	$\sigma_\eta^2 = 0$	$\sigma_\eta^2 \neq 0$
Die in den drei Zeilen „Ergebnis“ stehenden Angaben sind wie folgt zu interpretieren: $\sigma_i^2 = 0$: diese Varianz wird als nicht von Null verschieden angesehen. $\sigma_i^2 \neq 0$: diese Varianz wird als von Null verschieden betrachtet.		

σ_η^2 wird geschätzt durch

$$\hat{\sigma}_\eta^2 = \frac{s_S^2 - s_{ZS}^2}{n_1} \quad (33)$$

σ_ξ^2 wird geschätzt durch

$$\hat{\sigma}_\xi^2 = s_{ZS}^2 \quad (34)$$

Der Test für den Einfluß der Wechselwirkung nach Tabelle 3 entfällt. Der Test für den Zeileneinfluß und der Test für den Spalteneinfluß werden nach Tabelle 3 durchgeführt.

Falls die Hypothesen $\sigma_\xi^2 = 0$ und $\sigma_\eta^2 = 0$ nicht verworfen werden, ist der Vertrauensbereich des Mittelwertes der Gesamtheit μ zum Vertrauensniveau $1 - \alpha$:

$$\bar{x}_{..} - t_{f_{ZS}; 1-\alpha/2} \cdot \frac{s_{ZS}}{\sqrt{n_1 \cdot n_2}} \leq \mu \leq \bar{x}_{..} + t_{f_{ZS}; 1-\alpha/2} \cdot \frac{s_{ZS}}{\sqrt{n_1 \cdot n_2}} \quad (35)$$

Falls die obengenannten Hypothesen verworfen werden, gilt näherungsweise:

$$\begin{aligned} \bar{x}_{..} - t_{f; 1-\alpha/2} \cdot \sqrt{\frac{\hat{\sigma}_\xi^2}{n_1} + \frac{\hat{\sigma}_\eta^2}{n_2} + \frac{\hat{\sigma}_\epsilon^2}{n_1 \cdot n_2}} &\leq \mu \\ &\leq \bar{x}_{..} + t_{f; 1-\alpha/2} \cdot \sqrt{\frac{\hat{\sigma}_\xi^2}{n_1} + \frac{\hat{\sigma}_\eta^2}{n_2} + \frac{\hat{\sigma}_\epsilon^2}{n_1 \cdot n_2}} \end{aligned} \quad (36)$$

Der Freiheitsgrad für den Tabellenwert der t -Verteilung ist der kleinere der beiden Werte f_Z oder f_S .

Bezüglich der übrigen Vertrauensbereiche wird auf das Schrifttum verwiesen. [1], [2]

5 Stichprobenaufbau beim Modell mit Wechselwirkung

Beim Modell mit Wechselwirkung ist $\sigma_\alpha^2 \neq 0$.

Eine vorgegebene halbe absolute Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit μ , d. h. ein Vertrauensbereich der Form $\bar{x}_{..} - W \leq \mu \leq \bar{x}_{..} + W$ wird bei außerdem vorgeschriebenem Vertrauensniveau $1 - \alpha$ erreicht, wenn die unten berechneten Anzahlen n_1, n_2 und n_3 von Einheiten untersucht werden.

Aus dem vorgegebenen Wert W berechnet man zum Vertrauensniveau $1 - \alpha$ die Varianz σ_0^2 nach

$$\sigma_0^2 = \left(\frac{W}{u_{1-\alpha/2}} \right)^2 \quad (37)$$

wobei $u_{1-\alpha/2}$ der Tabellenwert der standardisierten Normalverteilung bei zweiseitiger Abgrenzung ist.

5.1 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit; varianzbestimmte Stichprobe

Die Varianz s_G^2 für den Mittelwert der Stichprobe $\bar{x}_{..}$ ist

$$s_G^2 = \frac{\hat{\sigma}_\xi^2}{n_1} + \frac{\hat{\sigma}_\eta^2}{n_2} + \frac{\hat{\sigma}_\alpha^2}{n_1 \cdot n_2} + \frac{\hat{\sigma}_\epsilon^2}{n_1 \cdot n_2 \cdot n_3} \quad (38)$$

Die Forderung lautet also $s_G^2 \leq \sigma_0^2$. Damit erhält man zunächst folgende im allgemeinen nicht-ganzzahligen Werte:

Fall a:

$$\begin{aligned} n'_{1a} &= \frac{\hat{\sigma}_\xi^2}{\sigma_0^2} & n'_{2a} &= \frac{n_{1a+} \cdot \hat{\sigma}_\eta^2 + \hat{\sigma}_\alpha^2}{n_{1a+} \cdot \sigma_0^2 - \hat{\sigma}_\xi^2} \\ n'_{3a} &= \frac{\hat{\sigma}_\epsilon^2}{n_{1a+} \cdot n_{2a+} \cdot \sigma_0^2 - n_{2a+} \cdot \hat{\sigma}_\xi^2 - n_{1a+} \cdot \hat{\sigma}_\eta^2 - \hat{\sigma}_\alpha^2} \end{aligned} \quad (39)$$

Dabei sind n_{1a+} und n_{2a+} die zur nächst größeren ganzen Zahl aufgerundeten Werte n'_{1a} und n'_{2a} oder die um 1 größeren Werte, falls n'_{1a} und/oder n'_{2a} ganzzahlig sein sollten. n'_{3a} wird zu n_{3a+} aufgerundet oder bleibt unverändert, falls n'_{3a} ganzzahlig ist.

Weitere Ausführungen zu dieser Rundungsempfehlung sind in Abschnitt 7 zusammengestellt.

Der Wert n_{1a+} darf nicht unterschritten werden, wenn die Bedingung $s_G^2 \leq \sigma_0^2$ erfüllt werden soll; zu diesem n_{1a+} gehören die Werte n_{2a+} und n_{3a+} .

Fall b:

$$\begin{aligned} n'_{2b} &= \frac{\hat{\sigma}_\eta^2}{\sigma_0^2} & n'_{1b} &= \frac{n_{2b+} \cdot \hat{\sigma}_\xi^2 + \hat{\sigma}_\alpha^2}{n_{2b+} \cdot \sigma_0^2 - \hat{\sigma}_\eta^2} \\ n'_{3b} &= \frac{\hat{\sigma}_\epsilon^2}{n_{1b+} \cdot n_{2b+} \cdot \sigma_0^2 - n_{2b+} \cdot \hat{\sigma}_\xi^2 - n_{1b+} \cdot \hat{\sigma}_\eta^2 - \hat{\sigma}_\alpha^2} \end{aligned} \quad (40)$$

Die Werte n'_{1b} , n'_{2b} und n'_{3b} werden analog zu Fall a aufgerundet. Der Wert n_{2b+} darf nicht unterschritten werden, wenn die Bedingung $s_G^2 \leq \sigma_0^2$ erfüllt werden soll; zu diesem n_{2b+} gehören die Werte n_{1b+} und n_{3b+} .

Setzt man $n_{3a} = 1$ im Fall a, dann erhält man zu dem Wert n_{1a+} das Wertetripel $(n_{1a+}; n_{2a+}; 1)$, wobei n_{2a+} der Wert ist, der nicht überschritten zu werden braucht, um die Bedingung $s_G^2 \leq \sigma_0^2$ einzuhalten. Zunächst berechnet man:

$$n''_{2a} = \frac{\hat{\sigma}_\eta^2 + \frac{\hat{\sigma}_\alpha^2}{n_{1a+}} + \frac{\hat{\sigma}_\epsilon^2}{n_{1a+}}}{\sigma_0^2 - \frac{\hat{\sigma}_\xi^2}{n_{1a+}}} = \frac{n_{1a+} \cdot \hat{\sigma}_\eta^2 + \hat{\sigma}_\alpha^2 + \hat{\sigma}_\epsilon^2}{n_{1a+} \cdot \sigma_0^2 - \hat{\sigma}_\xi^2} \quad (41)$$

Diesen Wert rundet man zu n_{2a++} auf.

Für $n_{3b} = 1$ erhält man im Fall b zu dem Wert n_{2b+} das Wertetripel $(n_{1b++}; n_{2b+}; 1)$, wobei n_{1b++} der Wert ist, der nicht überschritten zu werden braucht, um die Bedingung $s_G^2 \leq \sigma_0^2$ einzuhalten. Zunächst berechnet man:

$$n''_{1b} = \frac{\hat{\sigma}_\xi^2 + \frac{\hat{\sigma}_\alpha^2}{n_{2b+}} + \frac{\hat{\sigma}_\epsilon^2}{n_{2b+}}}{\sigma_0^2 - \frac{\hat{\sigma}_\eta^2}{n_{2b+}}} = \frac{n_{2b+} \cdot \hat{\sigma}_\xi^2 + \hat{\sigma}_\alpha^2 + \hat{\sigma}_\epsilon^2}{n_{2b+} \cdot \sigma_0^2 - \hat{\sigma}_\eta^2} \quad (42)$$

Diesen Wert rundet man zu n_{1b++} auf.

Die Bereiche für die Anzahl der Einheiten in den Zeilen, Spalten und Zellen sind damit:

$$\begin{aligned} 1 &\leq n_{1a+} \leq n_1 \leq n_{1b++} \\ 1 &\leq n_{2b+} \leq n_2 \leq n_{2a++} \\ 1 &\leq n_3 \leq \max(n_{3a+}; n_{3b+}) \end{aligned} \quad (43)$$

Die Obergrenze für n_3 wird durch den größeren der beiden Werte n_{3a+} und n_{3b+} gegeben. In beiden Fällen a und b erfüllt die Stichprobe mit den Werten n_{1+} , n_{2+} und n_{3+} die geforderte Bedingung; dabei ist n_{1+} die kleinstmögliche Anzahl für die Aufteilung nach dem Zeilengesichtspunkt, mit der die Bedingung $s_G^2 \leq \sigma_0^2$ erfüllbar ist. Zu diesem n_{1+} gehören mindestens die Anzahlen n_{2+} und n_{3+} . Der Stich-

probenumfang $n_{1+} \cdot n_{2+} \cdot n_{3+}$ ist jedoch trotz des Mindestwertes n_{1+} nicht der kleinstmögliche. Man kann ihn verkleinern, wenn man bei unverändertem n_{1+} den kleineren der beiden Werte n_{2+} und n_{3+} etwas vergrößert und den anderen dieser Werte deutlich verkleinert.

Den gesamten Stichprobenumfang $n_{1+} \cdot n_{2+} \cdot n_{3+}$ kann man weiterhin durch Vergrößern von n_{1+} auf verschiedene Werte n_1 ($n_1 > n_{1+}$) verkleinern, wenn man die Anzahlen n_{2+} und n_{3+} beibehält oder einzeln oder beide **verkleinert** ($n_2 \leq n_{2+}$ und $n_3 \leq n_{3+}$).

Den kleinsten Stichprobenumfang erhält man, wenn man $n_3 = 1$ setzt und gleichzeitig den kleinstmöglichen Wert für n_1 oder n_2 wählt oder für n_1 oder n_2 einen Wert wählt, der nur wenig größer als der kleinstmögliche ist. Der kleinstmögliche Stichprobenumfang ist im allgemeinen jedoch nicht der kostengünstigste.

Beginnend mit $(n_{1a+}; n_{2a+}; n_{3a+})$ oder $(n_{2b+}; n_{1b+}; n_{3b+})$ werden die Wertetripel innerhalb der Bereiche nach Gleichung (43) variiert. Dabei werden der Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ sowie alle Wertetripel notiert, für die s_G^2 nur wenig kleiner als σ_0^2 ist.

Nach fallweise verschiedenen Gesichtspunkten wählt man aus den möglichen Wertetripeln ein Tripel aus.

Falls einer der zunächst als nicht-ganzzahlig berechneten Werte n'_1 , n'_2 und n'_3 kleiner als 1 ist, ist nach Abschnitt 9 vorzugehen.

Beispiel 1:

Bei einer Untersuchung haben sich folgende Schätzwerte für die Varianzen ergeben:

$$\hat{\sigma}_\xi^2 = 64 \quad \hat{\sigma}_\eta^2 = 36 \quad \hat{\sigma}_\alpha^2 = 49 \quad \hat{\sigma}_\epsilon^2 = 225$$

Auf dem Vertrauensniveau $1 - \alpha$ soll die Varianz $\sigma_0^2 = 30$ nicht überschritten werden. (Für dieses und die folgenden grundsätzlichen Beispiele bleibt die Einheit der Varianzen außer Betracht.)

Nach den in Abschnitt 5.1 angegebenen Gleichungen (39), (40), (41) und (42) berechnet man:

$$\begin{aligned} n'_{1a} &= 2,13 \text{ aufgerundet auf } n_{1a+} = 3 \\ n'_{2a} &= 6,04 \text{ aufgerundet auf } n_{2a+} = 7 \\ n'_{3a} &= 9,00 \text{ also } n_{3a+} = 9 \\ n'_{2b} &= 1,20 \text{ aufgerundet auf } n_{2b+} = 2 \\ n'_{1b} &= 7,38 \text{ aufgerundet auf } n_{1b+} = 8 \\ n'_{3b} &= 15,00 \text{ also } n_{3b+} = 15 \\ n''_{2a} &= 14,69 \text{ aufgerundet auf } n_{2a++} = 15 \\ n''_{1b} &= 16,75 \text{ aufgerundet auf } n_{1b++} = 17 \end{aligned}$$

Die Bereiche, in denen n_1 , n_2 und n_3 variiert werden können, sind damit in diesem Falle:

$$3 \leq n_1 \leq 17 \quad 2 \leq n_2 \leq 15 \quad 1 \leq n_3 \leq 15$$

In der Tabelle 4 sind **sämtliche** günstigen Kombinationen $(n_1; n_2; n_3)$ für dieses Beispiel zusammengestellt, für die s_G^2 nur wenig kleiner als σ_0^2 ist. In diesem Falle wurde die Berechnung mit dem kleinsten Wert für n_{2+} begonnen.

Diese Kombinationen erhält man auf folgende Weise: Man beginnt mit dem Wertetripel $(n_{1+}; n_{2+}; n_{3+}) = (8; 2; 15)$. $n_{2+} = 2$ ist die kleinstmögliche Anzahl der Spalten, mit der man die Bedingung $s_G^2 \leq \sigma_0^2$ einhalten kann; $n_1 = 17$ und $n_3 = 15$ sind die größten Werte, die man zur Einhaltung der Bedingung $s_G^2 \leq \sigma_0^2$ nicht zu überschreiten braucht. Weitere Wertetripel, für die s_G^2 möglichst wenig kleiner als σ_0^2 ist, erhält man, wenn man n_1 um mindestens 1 vergrößert und **gleichzeitig** n_3 um mindestens 1 verkleinert. Für $n_2 = 2$ ergibt sich als nächstes Wertetripel $(9; 2; 6)$; dabei wurde $n_1 = 9$ als der um 1 vergrößerte vorhergehende Wert angesetzt, und n_3 der um 9 verkleinerte Wert.

Tabelle 4. Wertetripel ($n_1; n_2; n_3$) und Varianz für den Mittelwert der Stichprobe s_G^2 für Beispiel 1

n_1	n_2	n_3	$n_1 \cdot n_2 \cdot n_3$	s_G^2
8	2	15	240	30,00
9		6	108	29,92
10		4	80	29,66
11		3	66	29,45
13		2	52	29,13
17		1	34	29,82
5	3	8	120	29,94
6		3	54	29,56
7		2	42	28,83
9		1	27	29,26
4	4	8	128	29,82
5		2	40	29,88
7		1	28	27,93
4	5	3	60	29,40
5		2	50	26,46
6		1	30	27,00
4	6	2	48	28,73
5		1	30	27,93
3	7	9	189	30,00
4		2	56	26,91
5		1	35	25,77
3	8	5	120	29,75
4		1	32	29,06
3	9	3	81	29,93
4		1	36	27,61
3	11	2	66	29,50
4		1	44	25,50
3	15	1	45	29,82

wurde $n'_{3b} = 5,77$ berechnet, also aufgerundet $n_3 = 6$. Genau so verfährt man für $n_2 = 2$ mit den für n_1 nacheinander festgesetzten Werten 10 und 11. Für $n_1 = 12$ ist $n'_{3b} = 2,03$, also aufgerundet $n_3 = 3$. Dieses Tripel (12; 2; 3) ergibt gegenüber dem Tripel (11; 2; 3) eine Varianz s_G^2 , die merklich kleiner als 30 ist, weil mit dem um 1 vergrößerten Wert n_1 nicht gleichzeitig n_3 um 1 kleiner wird. Deshalb wird das Wertetripel (12; 2; 3) weggelassen. Erst das Tripel (13; 2; 2) ergibt wieder einen Wert s_G^2 , der etwas kleiner als 30 ist; gegenüber diesem Tripel werden (14; 2; 2), (15; 2; 2) und (16; 2; 2) weggelassen. Bei (17; 2; 1) erhält man die letzte günstige Kombination für $n_2 = 2$, weil n_3 nicht kleiner als 1 werden kann.

Weiter setzt man $n_2 = 3$ und berechnet aus der Gleichung (40) $n'_{1b} = 4,46$, also aufgerundet $n_1 = 5$; es folgt $n'_{3b} = 7,76$, das zu $n_3 = 8$ aufgerundet wird. Mit dem Wertetripel (5; 3; 8) beginnend rechnet man weiter wie oben angegeben. Beim Wertetripel (9; 3; 1) wird $n_3 = 1$, die weitere Vergrößerung von n_1 über $n_1 = 9$ hinaus ist ungünstig.

Bei $n_{2++} = 15$ erhält man das letzte günstige Wertetripel (3; 15; 1). Die Vergrößerung von n_2 über $n_2 = 15$ hinaus ist ungünstig.

Den kleinstmöglichen günstigen Stichprobenumfang $n_1 \cdot n_2 \cdot n_3 = 27$ erhält man mit dem Tripel

(9; 3; 1). Nach von Fall zu Fall unterschiedlichen Gesichtspunkten wählt man eines dieser Wertetripel oder eines der in der Nähe dieser Tripel liegenden Wertetripel mit **größerem** Gesamtumfang $n_1 \cdot n_2 \cdot n_3$ für die zu ziehende Stichprobe aus. Wünscht man z. B. $n_1 = 10$ und $n_3 = 5$, dann muß $n_2 \geq 2$ sein. Für $n_2 = 1$ wird in diesem Falle $s_G^2 > \sigma_0^2$. Wie dieses Beispiel zeigt, kann man durch mäßiges Vergrößern von n_2 über n_{2+} hinaus die Werte für n_1 und n_3 verändern, wobei sich in jedem Falle der Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ drastisch reduziert.

5.2 Stichprobenaufbau bei vorgegebener Weite des Vertrauensbereiches für den Mittelwert der Gesamtheit und minimalen Gesamtkosten; varianzbestimmte kostenminimale Stichprobe

In diesem Falle ist der Aufbau der varianzbestimmten kostenminimalen Stichprobe nur mit Hilfe eines aufwendigen Iterationsverfahrens oder näherungsweise zu berechnen. Die für die Näherungslösung notwendigen Voraussetzungen sind jedoch nicht einfach zu erfüllen. Deshalb soll das Problem in dieser Norm nicht behandelt werden. [3]

5.3 Stichprobenaufbau bei vorgegebenen Gesamtkosten; kostenbestimmte Stichprobe

Ist an Stelle der Varianz σ_0^2 ein oberer Grenzwert K_0 für die Gesamtkosten der Untersuchung vorgegeben, der nicht überschritten werden soll, kann man ausgehend von der in diesem Fall gültigen Gleichung für die Gesamtkosten der Untersuchung

$$K_G = n_1 \cdot c_\xi + n_2 \cdot c_\eta + n_1 \cdot n_2 \cdot n_3 \cdot c_\epsilon \quad (44)$$

ebenfalls den Stichprobenaufbau berechnen. Hierbei sind c_ξ und c_η die Kosten (in willkürlichen Einheiten) für die Bereitstellung einer Einheit in Zeilen- bzw. Spaltenrichtung und c_ϵ für die Bereitstellung einer Endprobe und deren Messung. Diese Kostengleichung ist von der Wechselwirkung unabhängig. Mit der Bedingung $K_G \leq K_0$ erhält man zunächst folgende im allgemeinen nicht-ganzzahlige Werte für die Anzahl der Zeilen n_1 und der Spalten n_2 sowie die Anzahl n_3 der Einzelwerte in jeder Zelle:

$$\begin{aligned} n'_1 &= \frac{K_0 - c_\eta}{c_\xi + c_\epsilon} & n'_2 &= \frac{K_0 - n_{1-} \cdot c_\xi}{c_\eta + n_{1-} \cdot c_\epsilon} \\ n'_3 &= \frac{K_0 - n_{1-} \cdot c_\xi - n_{2-} \cdot c_\eta}{n_{1-} \cdot n_{2-} \cdot c_\epsilon} \end{aligned} \quad (45)$$

Dabei sind n_{1-} und n_{2-} die zur nächsten ganzen Zahl abgerundeten Werte n'_1 und n'_2 . Falls n'_1 und n'_2 ganzzahlig sind, bleiben sie **unverändert**. n'_3 wird zur nächsten ganzen Zahl n_{3-} abgerundet, sofern es nicht ganzzahlig ist.

Die Stichprobe mit dem kleinsten Umfang ist diejenige, für die $n_1 = n_2 = n_3 = 1$ ist. Deshalb kann der vorgegebene Wert K_0 nicht beliebig klein gemacht werden, sondern er muß die Bedingung

$$K_0 \geq c_\xi + c_\eta + c_\epsilon \quad (46)$$

erfüllen.

Die hier angegebenen Beziehungen für n'_i sind bei $c_\xi > c_\eta > c_\epsilon$ rechenstechnisch günstiger als die fünf Sätze von anderen möglichen Gleichungen für die Werte n'_i . Alle sechs Sätze von Beziehungen ergeben dieselben Kombinationen für den Stichprobenumfang.

Die Obergrenze der Anzahl der Zeilen, der Spalten und der Einzelwerte innerhalb einer Zelle, die für $K_G \leq K_0$ nicht überschritten werden darf, ist:

$$n''_1 = n'_1 = \frac{K_0 - c_\eta}{c_\xi + c_\epsilon} \quad n''_2 = \frac{K_0 - c_\xi}{c_\eta + c_\epsilon} \quad n''_3 = \frac{K_0 - c_\xi - c_\eta}{c_\epsilon} \quad (47)$$

Diese Werte werden zu n_{1--} , n_{2--} und n_{3--} **abgerundet**. Die Bereiche für die Anzahl der Einheiten in den Zeilen, Spalten und Zellen sind damit:

$$1 \leq n_1 \leq n_{1--} = n_{1-} \quad 1 \leq n_2 \leq n_{2--} \quad 1 \leq n_3 \leq n_{3--} \quad (48)$$

Die günstigen Wertetripel $(n_1; n_2; n_3)$ für den Stichprobenumfang erhält man analog zu dem in Abschnitt 5.1 dargestellten Verfahren. Dabei sind jedoch folgende Unterschiede zu beachten:

- Die zunächst berechneten Werte n'_i sind **abzurunden**.
- Falls ein n'_i -Wert ganzzahlig ist, bleibt er **unverändert**.
- Den kleinstmöglichen Stichprobenumfang erhält man für $n_1 = n_2 = n_3 = 1$.
- Es gibt für die Anzahl der Zeilen, Spalten und der Einzelwerte in einer Zelle die angegebenen oberen Grenzwerte n''_i , die nicht überschritten werden dürfen, wenn die Bedingung $K_G \leq K_0$ eingehalten werden soll.
- Ausgehend von den abgerundeten Werten n_{i-} , die aus n'_i erhalten werden, kann man den Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ weiter **vergrößern**, wenn man n_1 verkleinert von n_{1-} bis hinab zu $n_1 = 1$ und dabei n_2 und n_3 einzeln oder beide vergrößert.
- Den größten Stichprobenumfang erhält man für $n_1 = n_2 = 1$ und n_{3--} als abgerundeten Wert n''_3 . Dieser größtmögliche Stichprobenumfang ergibt im allgemeinen nicht die kleinste erreichbare Varianz s_G^2 .

Beginnend mit $(n_{1-}; n_{2-}; n_{3-})$ werden die Anzahlen der Einheiten in den Zeilen, Spalten und Zellen innerhalb der Bereiche nach Gleichung (48) variiert. Dabei werden der Stichprobenumfang $n_1 \cdot n_2 \cdot n_3$ sowie alle Wertetripel notiert, für die K_G nur wenig kleiner als K_0 ist.

Nach fallweise verschiedenen Gesichtspunkten wählt man aus den möglichen Wertetripeln ein Tripel aus.

Falls einer der zunächst als nicht-ganzzahlig berechneten Werte n'_1 , n'_2 und n'_3 kleiner als 2 ist, ist nach Abschnitt 9 vorzugehen.

Zur Berechnung der Zahlentripel, mit denen die Bedingung $K_G \leq K_0$ erfüllbar ist, benötigt man nur die vorgegebenen Gesamtkosten K_0 und die Einzelkosten c_i .

Beispiel 2:

Für eine Untersuchung gelten folgende Einzelkosten (in willkürlichen Einheiten):

$$c_\xi = 25 \quad c_\eta = 5 \quad c_\varepsilon = 1$$

Die Gesamtkosten der Untersuchung $K_0 = 170$ sollen nicht überschritten werden. Nach den in Abschnitt 5.3 angegebenen Gleichungen (45) und (47) berechnet man:

$$\begin{aligned} n'_1 &= 6,35 \text{ abgerundet auf } n_{1-} = 6 \\ n'_2 &= 1,82 \text{ abgerundet auf } n_{2-} = 1 \\ n'_3 &= 2,50 \text{ abgerundet auf } n_{3-} = 2 \\ n''_2 &= 24,17 \text{ abgerundet auf } n_{2--} = 24 \\ n''_3 &= 140,00 \text{ also } n_{3--} = 140 \end{aligned}$$

Die Bereiche, in denen n_1 , n_2 und n_3 variiert werden können, sind damit in diesem Falle:

$$1 \leq n_1 \leq 6 \quad 1 \leq n_2 \leq 24 \quad 1 \leq n_3 \leq 140$$

In der Tabelle 5 sind sämtliche günstigen Kombinationen des Stichprobenumfanges $n_1 \cdot n_2 \cdot n_3$ angegeben sowie die dazugehörigen Gesamtkosten K_G .

Diese Kombinationen werden analog zu dem in Beispiel 1 angegebenen Verfahren unter Beachtung

der in Abschnitt 5.3 aufgeführten Einzelheiten berechnet.

Nach von Fall zu Fall unterschiedlichen Gesichtspunkten wählt man eines der in der Tabelle angegebenen Wertetripel oder eines der in der Nähe dieser Werte liegenden Tripel mit **kleinerem** Gesamtumfang $n_1 \cdot n_2 \cdot n_3$ für die zu ziehende Stichprobe aus.

Tabelle 5. Wertetripel $(n_1; n_2; n_3)$ und Gesamtkosten K_G für Beispiel 2

n_1	n_2	n_3	$n_1 \cdot n_2 \cdot n_3$	K_G
6	1	2	12	167
5	4	1	20	165
	3	2	30	170
	2	3	30	165
	1	8	40	170
4	7	1	28	163
	5	2	40	165
	4	3	48	168
	3	4	48	163
	2	7	56	166
	1	16	64	169
3	11	1	33	163
	8	2	48	163
	6	3	54	159
	5	4	60	160
	4	6	72	167
	3	8	72	162
	2	14	84	169
	1	30	90	170
2	17	1	34	169
	13	2	52	167
	10	3	60	160
	9	4	72	167
	8	5	80	170
	7	6	84	169
	6	7	84	164
	5	9	90	165
	4	12	96	166
	3	17	102	167
	2	27	108	168
	1	57	114	169
1	24	1	24	169
	20	2	40	165
	18	3	54	169
	16	4	64	169
	14	5	70	165
	13	6	78	168
	12	7	84	169
	11	8	88	168
	10	9	90	165
	9	11	99	169
	8	13	104	169
	7	15	105	165
	6	19	114	169
	5	24	120	170
	4	31	124	169
	3	43	129	169
	2	67	134	169
	1	140	140	170